STAIRS 2006
Pavlos Peppas, Anna Perini and Loris Penserini
IOS Press, 2006

Argumentation Semantics for Temporal Defeasible Logic[1]

Régis Riveret [a,2]x, Guido Governatori [b], and Antonino Rotolo [a]
[a] *CIRSFID, University of Bologna*
[b] *School of ITEE, University of Queensland*

Keywords. Defeasible Logic, Argumentation Semantics, Temporal Reasoning

Temporal Defeasible Logic extends Defeasible Logic (DL) [1] to deal with temporal aspects. This extension proved useful in modelling temporalised normative positions [3] and retroactive rules, which permit to obtain conclusions holding at a time instant that precedes the time of application of the same rules [4]. Time is added in two ways. First, a temporalised literal is a pair $l : t$ where l is a literal and t is an instant of time belonging to a discrete totally ordered set of instants of time $\mathscr{T} = \{t_1, t_2, \dots\}$. Intuitively, the meaning of a temporalised literal $l : t$ is that l holds at time t. Second, rules are partitioned in persistent and transient rules according to whether the consequent persists until an interrupting event occurs or is co-occurrent with the premises. Hence, $D = (\mathscr{T}, F, R^p, R^c, \succ)$ is a temporal defeasible theory, where $\mathscr{T}$ is the set of instants, F a set of facts, $\succ$ a superiority relation over rules, and R^p and R^t the sets of persistent and transient rules. Given a rule $r \in R^p$ such as $a : t \Rightarrow_p b : t'$, if r is applicable, we can derive b holding at t' and at any $t' + n$, until we can block this inference, for example, by deriving $\neg b$, at a certain time $t' + m$; given an $r' \in R^t$ such as $a : t \Rightarrow_t b : t$, we derive b at any t only if a holds as well at t. Proof tags of DL must be duplicated: $\pm\Delta_p l : t$ and $\pm\partial_p l : t$ mean, respectively, that $l : t$ is/is not definitely and defeasibly provable persistently; $\pm\Delta_t l : t$ and $\pm\partial_t l : t$ mean, respectively, that $l : t$ is/is not definitely and defeasibly provable transiently.

On the other hand, DL can also be interpreted in terms of interacting arguments, giving for it an argumentation semantics [2]. Argumentation systems are of particular interest in AI & Law, where notions like argument and counter-argument are very common. For example, a recent development of such semantics is represented by argumentation and mediation systems which assist the users in expressing and organising their arguments, in assessing their impact on controversial legal issues or in building up an effective interaction in dialectical contexts [5]. So far, the logic has been only formalised in a proof-theoretic setting in which arguments play no role. Our purpose is to provide an argumentation semantics for temporal DL. Note that we can dispense with the superi-

[1] The full version of this paper is available at `http://eprint.uq.edu.au/archive/00003954/01/Stairs06Long.pdf`.

[2] Corresponding Author: Régis Riveret, CIRSFID, University of Bologna, 40121 Bologna, Italy; E-mail: rriveret@cirsfid.unibo.it. The first and third authors were supported by the European project for Standardized Transparant Representations in order to Extend Legal Accessibility (ESTRELLA, IST-4-027655); the second author was supported by the Australia Research Council under Discovery Project No. DP0558854 on "A Formal Approach to Resource Allocation in Service Oriented Marketplaces".

ority relation of standard DL, as we can define a modular transformation similar to that given in [1] that enables to empty the superiority relation.

In line with [2], an argument for a temporalised literal $p : t$ is a proof tree (or monotonic derivation) for that temporalised literal in DL. Nodes are temporalised literals, arcs connecting nodes correspond to rules. We introduce a new type of connection, represented by "dashed arcs", which are meant to connect a permanent literal to its successors in time. Types of arguments are distinguished according to the rules used: supportive arguments are finite arguments in which no defeater is used and strict arguments use strict rules only, permanent arguments are arguments in which the ending arc either is a dashed arc or corresponds to a permanent rule while in transient arguments the ending arc corresponds to a transient rule.

The notion of attack between arguments is thus defined: a set of arguments S attacks a defeasible argument B if there is an argument A in S that attacks B such that (i) $a : t_a$ is a conclusion of A and (ii) $a : t_a$ is not a conclusion of a dashed arc, and (iii) $b : t_b$ is a conclusion of B, and a is on conflict with b (typically a is the complement of b), and either $t_a = t_b$, or it exists a dashed arc in B with premise $b : t_{b'}$ and conclusion $b : t_b$, and $t_{b'} < t_a \leq t_b$. This definition of attack allows us to reuse the standard definition as given in [2] of arguments being undercut and arguments being acceptable. Based on these concepts we proceed to define justified arguments, i.e., arguments that resist any refutation. Accordingly, a literal $p : t$ is defined as transiently/permanently justified if it is the conclusion of a supportive and transient/permanent argument in the set of justified arguments $Jargs_D$ of a theory D. That a literal $p : t$ is justified means that it is provable ($+\partial$). However, DL permits to express when a conclusion is not provable ($-\partial$). This last notion is captured by assigning the status of rejected to arguments. Roughly, an argument is rejected if it has a rejected sub-argument or it cannot overcome an attack from a justified argument. More generally, a literal is transiently/permanently rejected if it is transiently/permanently rejected by $Jargs_D$. One of our results is that permanent and transient defeasible conclusions can be characterised as follows:

THEOREM 1 *Given a theory D and its set of justified arguments $Jargs_D$,*

- $D \vdash +\partial_p(p : t)$ *iff* $p : t$ *is permanently justified;*
- $D \vdash -\partial_p(p : t)$ *iff* $p : t$ *is permanently rejected by* $Jargs_D$*;*
- $D \vdash +\partial_t(p : t)$ *iff* $p : t$ *is transiently justified;*
- $D \vdash -\partial_t(p : t)$ *iff* $p : t$ *is transiently rejected by* $Jargs_D$*.*

Hence, Theorem 1 gives an argumentation semantics with ambiguity blocking to permanent and transient defeasible conclusions.

References

[1] G. Antoniou, D. Billington, G. Governatori, and M.J. Maher. Representation results for defeasible logic. *ACM Transactions on Computational Logic*, 2, pages 255–287, 2001.

[2] G. Governatori, M.J. Maher, D. Billington, and G. Antoniou. Argumentation semantics for defeasible logics. *Journal of Logic and Computation*, 14, pages 675–702, 2004.

[3] G. Governatori, A. Rotolo, and G. Sartor. Temporalised normative positions in defeasible logic. In *Proc. ICAIL05*. ACM, New York, 2005.

[4] G. Governatori, M. Palmirani, R. Riveret, A. Rotolo, and G. Sartor. Norm modifications in defeasible logic. In M. Moens, editor, *Proc. Jurix'05*. IOS Press, Amsterdam, 2005.

[5] B. Verheij. *Virtual Arguments: On the Design of Argument Assistants for Lawyers and Other Arguers*. T.M.C. Asser Press, The Hague, 2005.